



ANNUAL REVIEWS **Further**

Click [here](#) to view this article's
online features:

- Download figures as PPT slides
- Navigate linked references
- Download citations
- Explore related articles
- Search keywords

Adversarial Allegiance among Expert Witnesses

Daniel C. Murrie¹ and Marcus T. Boccaccini²

¹Institute of Law, Psychiatry, and Public Policy, University of Virginia, Charlottesville,
Virginia 22903; email: Murrie@Virginia.edu

²Psychology Department, Sam Houston State University, Huntsville, Texas 77341

Annu. Rev. Law Soc. Sci. 2015. 11:37–55

First published online as a Review in Advance on
August 20, 2015

The *Annual Review of Law and Social Science* is
online at lawsocsci.annualreviews.org

This article's doi:
[10.1146/annurev-lawsocsci-120814-121714](https://doi.org/10.1146/annurev-lawsocsci-120814-121714)

Copyright © 2015 by Annual Reviews.
All rights reserved

Keywords

bias, allegiance, adversarial, forensic science, forensic assessment, expert
witness, expert testimony

Abstract

Can experts remain objective and accurate when one particular side in adversarial legal proceedings retains them? Despite long-standing concerns from the legal system and the general public, research has only recently explored whether experts can provide opinions unbiased by the side that retained them. This review addresses some of the factors that may cause disagreements between opposing experts. After summarizing recent field and experimental studies on mental health evaluations by forensic experts, we conclude that working for one side in an adversarial case causes some experts' opinions to drift toward the party retaining their services, even on ostensibly objective instruments and procedures. We call this process adversarial allegiance. The mechanisms that underlie adversarial allegiance among forensic experts are likely similar to the unconscious heuristics and cognitive biases that compromise judgment in a variety of other settings, but these will require further study to understand, and ultimately reduce, adversarial allegiance.

INTRODUCTION

Can experts who are retained by one side in adversarial legal proceedings offer the court genuinely objective findings and expert opinions? Or are these experts inevitably biased by the adversarial arrangements in which they work? In other words, does what we term “adversarial allegiance” to the retaining party lead experts to drift from strictly objective findings toward findings that better support the party that retained them?

Questions about adversarial allegiance are long-standing and widespread. Expert witnesses—particularly those from medicine and the behavioral sciences—have long been tolerated with skepticism and sometimes characterized as hired guns who can reach any conclusion to support the party that retained them. Whereas broad concerns about expert bias have remained commonplace among legal professionals and the general public, there has been a striking lack of empirical research investigating the accuracy of these concerns, at least until quite recently.

Bias and other methodological problems compromise most types of forensic science procedures (Comm. Identif. Needs Forensic Sci. Community, Natl. Res. Counc. 2009), but to provide detailed illustration and discussion, our review focuses specifically on potential bias among forensic mental health experts (i.e., psychologists and psychiatrists). Recent research with forensic mental health experts provides an opportunity to begin disentangling the many factors that may lead to disagreements among opposing experts as well as to examine specifically the evidence for adversarial allegiance, the tendency for experts retained by one side in adversarial proceedings to drift from strictly objective conclusions toward conclusions that better support the retaining party.

This review first addresses the long-standing concerns about adversarial experts and the challenges that hinder efforts to study bias among experts. After reviewing the few historical field and experimental studies of expert bias, we describe a new generation of studies that used forensic assessment instruments as a metric to demonstrate and quantify the apparent influence of adversarial allegiance on forensic mental health experts. In light of compelling findings from these recent studies, we identify mechanisms that may contribute to adversarial allegiance and therefore warrant attention in future research and intervention efforts.

Long-Standing Concerns about Bias among Expert Witnesses

Legal scholars have lamented for more than a century that expert witnesses, particularly medical experts, are inevitably biased by the party that pays their fees (e.g., Foster 1897, Hand 1901, Wigmore 1923). Scholarly reviews have theorized about types of bias among expert witnesses (e.g., Bernstein 2008) and criticized experts for using “junk science” to support the retaining party’s case theory (e.g., Huber 1993). Beyond legal scholars, the judges and attorneys who observe expert testimony in the courtroom seem similarly skeptical. As early as 1889, one State Supreme Court justice lamented, “If there is any kind of testimony that is not only of no value, but even worse than that, it is . . . that of medical experts” (as cited in Mnookin 2007, p. 9). In a multistate survey conducted 100 years later, judges reported that bias was their primary frustration with expert witnesses (Shuman et al. 1994). More recently, in a large-scale survey of federal judges and attorneys, the most frequently cited concern about expert witnesses was that they “abandon objectivity and become advocates for the side that retained them” (Krafka et al. 2002, p. 328). One review of published cases even compiled the colorful terms—“hired guns,” “whores,” or some variant of “prostitute”—that legal professionals have used to describe mental health expert witnesses whom they perceived as biased (Mossman 1999). Beyond practicing jurists, the general public also appears distrustful of mental health experts, particularly well-paid experts (Boccaccini & Brodsky 2002, Cooper & Neuhaus 2000), perhaps assuming that paid work in an adversarial system ensures allegiance to the paying party.

Though apparently less cynical than legal scholars and the lay public about experts' motives, some authorities in the forensic mental health field have also acknowledged that the adversarial system may pull clinicians away from strict objectivity. Rogers (1987) warned clinicians about "forensic identification," which he described as the "unintentional process by which clinicians adopt the fact pattern or theory of the attorney with whom they have initial contact" (p. 151). Similarly, Brodsky (1991, pp. 8–10) identified a "pull to affiliate," which may lead forensic clinicians to gradually shape their opinions to align with the retaining party and become emotionally committed to the legal outcome that party desires. These historical warnings from forensic authorities seem to suggest that some form of bias (even if unconscious or unintentional) may be a significant risk in the field.

Also recognizing the potential for a pull to affiliate, the ethical guidelines for forensic mental health professionals direct them to strive for objective opinions, even in an adversarial system. For example, the original *Specialty Guidelines for Forensic Psychology* (Am. Psychol. Assoc. Comm. Ethical Guidel. Forensic Psychol. 1991) emphasized that psychologists who work in forensic contexts are bound by a "special responsibility for fairness and accuracy" and that "forensic psychologists take special care to avoid undue influence upon their methods, procedures, and products, such as might emanate from the party to a legal proceeding by financial compensation or other gains" (p. 661). Similarly, the *Ethics Guidelines for the Practice of Forensic Psychiatry* (Am. Acad. Psychiatry Law 2005, p. 3) warns that "the adversarial nature of most legal processes presents special hazards" and emphasizes that psychiatrists in legal proceedings "should adhere to the principle of honesty and strive for objectivity." In short, there has been widespread skepticism about forensic experts' ability to remain objective, but strong professional and ethical guidance encouraging them to do so. To what extent can experts actually resist the pull to affiliate and remain objective?

Challenges to Studying Expert Bias

Research, at least historically, has offered remarkably little data to shed light on these nagging questions about expert bias. In some respects, the lack of research data is puzzling, considering that the behavioral sciences have long appeared adept at identifying and studying allegiance bias within their own research (Blair et al. 2008; Boccaccini et al. in press; Luborsky et al. 1975, 1999) and other forms of unreliability within criminal justice procedures (e.g., Kassin et al. 2009, Wells et al. 1998). Indeed, we might expect that cognitive biases among experts in adversarial proceedings would be an irresistible topic for psychological study. But in other respects, this lack of research data is less surprising when we consider the practical challenges to studying adversarial allegiance or other biases among forensic experts.

Field studies. Perhaps the most straightforward way to study the possibility of expert bias is to conduct a field study by collecting opposing expert opinions from actual cases. For example, in what appears to be the first field study (and for decades, the only field study), researchers examined expert opinions relating to a lawsuit against a coal mining company following a 1972 disaster in West Virginia. The authors compared a set of psychiatric evaluations from plaintiff-retained and defense-retained experts for the same 42 litigants (Zusman & Simon 1983). They observed that plaintiff-retained psychiatrists arrived at findings that supported the plaintiffs, whereas defense-retained psychiatrists arrived at findings that supported the defendants.

Although the pattern of discrepant opinions between opposing experts in this West Virginia study is consistent with adversarial allegiance, such field research cannot rule out other plausible explanations. One possibility is that experts may have had access to different information, perhaps because attorneys provided them with different records or other sources, and therefore appropriately reached different conclusions on the basis of the data available to them. Another

possible explanation for the discrepant opinions involves the most important limitation of field studies of expert bias: Because experts are not randomly assigned to sides, the field study findings demonstrating differences in expert opinions may be a product of how attorneys select experts for their cases (selection effects), as opposed to the experts' decision making after they begin work on the cases (allegiance effects).

Indeed, research demonstrates substantial individual differences among experts in their tendencies to reach particular forensic opinions, a pattern of variability that may be exploited by savvy attorneys. For example, among forensic evaluators working within the same state, most concluded that 5–25% of the defendants they were ordered to evaluate for legal sanity met criteria for an insanity defense. But a few evaluators concluded that as many as 50% of defendants they evaluated were not guilty by reason of insanity, whereas others concluded that none were not guilty by reason of insanity (Murrie & Warren 2005). Likewise, individual evaluators differ in their patterns of finding defendants incompetent to stand trial (Murrie et al. 2008b) and their patterns of assigning scores on psychological assessment measures (Harris et al. 2014, McDermott et al. 2014), even when working for the same party and evaluating the same population of criminal offenders (Boccaccini et al. 2008, 2014). This individual variability may be due to evaluator training or sociopolitical values, or even evaluators' personality traits (Miller et al. 2011). A savvy defense attorney would almost certainly select an expert with a greater history of reaching opinions that support insanity, incompetence, or low violence risk; a savvy prosecuting attorney would almost certainly select an expert with a greater history of reaching opinions that support sanity, competence, or high violence risk. In short, good attorneys may capitalize on preexisting evaluator differences, and these attorney selection effects may explain some of the differences between opposing experts that we observe in the field.

Attorney selection as an explanation for field study findings presupposes that attorneys know about the leanings or case histories of potential experts, which appears common in some situations (e.g., Franklin 2012) but may be less likely in others. Even if they have little information about evaluators' leanings, attorneys with ample resources can still take advantage of evaluator differences by retaining multiple experts. If experts naturally vary in the conclusions they reach, either owing to preexisting proclivities or owing to the inherent subjectivity of some forensic judgments, retaining several experts increases the likelihood that at least one will come to a conclusion that supports the attorney's perspective. In most jurisdictions, the court is never exposed to opinions offered by "discarded experts" whom attorneys may retain for an initial evaluation but not subsequent reports or testimony (see *United States v. Alvarez* 1975). Only in some jurisdictions (see *United States ex rel. Edney v. Smith* 1976) are prosecutors allowed access to experts whom the defense has decided not to use. Therefore, another type of selection effect that could explain field study findings is that attorneys select only a subset of their retained experts—i.e., those who reached favorable conclusions—to present to the court. This means that courts (and therefore the public) are likely to see an overrepresentation of disagreeing experts and an underrepresentation of agreeing experts.

Of course, attorney selection effects are not the only selection effects that may contribute to apparent disagreements among experts. Evaluators may choose to work with one side of litigation but not the other, potentially skewing expert opinions to be more favorable to the retaining side. For example, evaluators who oppose the death penalty may select to work only for the defense in capital cases (Neal 2011). Perhaps in the early field study of opposing experts in a large lawsuit (Zusman & Simon 1983), experts more sympathetic toward victims and more likely to find trauma were more willing to work for the plaintiff, whereas those more skeptical of personal injury claims were more inclined to work for the defense. Such selection effects are almost impossible to eliminate, or even quantify, in a field study.

Experiments. The best way to control for any type of attorney or expert selection effect is to use an experimental design, with experts randomly assigned to one side or the other and no expert opinion excluded from analysis. However, conducting a realistic experiment with side of retention as the independent variable is exceedingly difficult. For example, in two early experiments, researchers provided to clinicians (or student clinicians) brief written vignettes to provide limited information about a sanity case and manipulate side of retention, then asked their opinion regarding legal sanity (Beckham et al. 1989, Otto 1989). Only one of these studies (Otto 1989) found evidence of adversarial allegiance, but neither came close to replicating the conditions that could promote adversarial allegiance in an actual forensic evaluation. The studies simply asked participants, via a brief written statement, to pretend that they were working for a particular side of the case. In most real cases, experts would have ongoing contact with an attorney, interview the defendant, seek and review a variety of collateral records, and understand that they will receive payment for their work—all opportunities for experts to feel the pull to affiliate and shift their opinions to be more favorable to the retaining side. Despite their limitations, these early vignette-based studies were important efforts to investigate a nearly hidden phenomenon. After all, it is hard to imagine a scenario in which courts or counsel would allow for a true random assignment of experts in real proceedings.

FORENSIC ASSESSMENT INSTRUMENTS PROVIDE A NEW METRIC TO STUDY BIAS

For decades, the one field study (Zusman & Simon 1983) and two vignette-based experiments (Beckham et al. 1989 and Otto 1989) remained the only empirical investigations addressing the possibility of adversarial allegiance among forensic mental health experts, leaving the field ill-equipped to answer public skepticism—or respond to the voluminous legal scholarship (e.g., Bernstein 2008; Diamond 1959; Gomez 2005; Mnookin 2007, 2008)—alleging pervasive expert bias. But during these same decades, the fields of forensic psychology and psychiatry advanced in many ways that may seem to reduce any problems of bias or poor methodology. For example, the fields developed a variety of specialized doctoral-level training programs, postdoctoral fellowships, continuing education programs for forensic professionals, and peer-reviewed scientific journals (Heilbrun & Brooks 2010, Otto & Heilbrun 2002). They produced more extensive ethical guidelines (Am. Psychol. Assoc. 2013) and have generally been progressing toward a clearer “standard of care” (Goldstein 2006). Contemporary forensic evaluation textbooks, guides for expert witnesses, and continuing education seminars all warn evaluators about the potential biasing effects of the adversarial system on evaluator opinions (e.g., Brodsky 2013, Commons et al. 2012, Otto et al. 2014). Perhaps the field’s increased attention to bias has served to inoculate evaluators from the potentially biasing effects of working within the adversarial system.

In addition to improved training and education for forensic mental health professionals, the past few decades have featured significant development in forensic mental health assessment: increased use of formal, standardized forensic assessment instruments to address legal questions surrounding adjudicative competence, risk of violence, risk of sexual recidivism, and similar issues (Grisso 2003, Otto & Douglas 2010, Rogers & Fiduccia 2015). Beginning in the late 1980s, scholars began developing a variety of tools aimed at structuring and reducing subjectivity in evaluator decision making. These instruments require evaluators to follow detailed rules to assign scores to the offender on a standardized set of items, which the evaluator then uses to assist with decision making. One goal of developing these measures was to reduce evaluators’ reliance on unstructured clinical judgment, which tends to be an unreliable, poor predictor of future outcomes (Ægisdóttir et al. 2006, Grove et al. 2000, Hanson & Morton-Bourgon 2009). As one authority

summarized, “The development of such specialized [forensic assessment] tools has been one of the most important influences in promoting evidence-based practice in this area. Using such a specialized tool is clearly consistent with best practice” (Heilbrun 2009, p. 123). Forensic evaluators increasingly administer these tests and describe test results when providing their opinions to the court. Some jurisdictions even mandate certain assessment instruments for particular evaluation or screening procedures (see Va. Code Ann. §37.2–903 or Tex. Health Saf. Code §841.023).

Structured forensic assessment instruments suggest a degree of rigor and reliability that should minimize subjectivity and bias. Consistent with this impression, research consistently documents that popular forensic assessment instruments demonstrate strong reliability in research contexts. In other words, different research assistants scoring the same offender in controlled settings tend to assign very similar scores. Forensic evaluators often cite the research describing an instrument’s strong reliability, typically reported in the instrument’s technical manual, to help justify presenting instrument scores as part of their expert opinions. Indeed, the strong reliability values associated with forensic assessment instruments may be one of the reasons that expert testimony based on these instruments is so often admitted as scientific evidence (see DeMatteo & Edens 2006, Vitacco et al. 2012).

Though the widespread use of structured forensic assessment instruments may seem likely to reduce any problematic bias among experts, this development also provides a new opportunity to study any problematic bias among experts. Specifically, the well-documented rater-agreement and error values for assessment instruments in research contexts can provide a point of comparison for the rater-agreement values on these instruments when administered in adversarial legal contexts. These well-documented rater-agreement values provide a clear expectation for the reliability we should observe in adversarial settings if adversarial allegiance has no effect on evaluators. Such reliability figures were not available for other types of forensic opinions—including opinions about emotional injury or legal sanity—in prior bias research (Beckham et al. 1989, Otto 1989, Zusman & Simon 1983). Furthermore, those types of expert opinions are more nebulous and less precisely quantifiable than assessment instrument scores, which more easily allow us to quantify the level of agreement among raters. In short, forensic assessment instruments allow researchers to easily quantify the level of disagreement among evaluators assigning scores in adversarial contexts and to compare that quantity to the known level of disagreement among evaluators assigning scores in nonadversarial contexts.

Thus, a general strategy for exploring bias among experts may be to examine real-world situations in which opposing experts administer the same forensic instruments to the same defendant. If adversarial contexts reveal poorer agreement, or more error, on instrument scores than is found in research from instrument manuals or well-designed studies, we may assume that the instrument or procedure simply has poorer reliability in field settings than in research settings. Such a finding would be unsurprising, given that highly controlled research studies of instrument reliability are likely to produce greater rater agreement than field situations where evaluators may vary widely on training, oversight, and available case information. However, if we were to find that the instrument has poorer agreement in the field and that the direction of disagreement appears systematically related to the forensic expert’s side of retention, then we may infer the poorer agreement is somehow attributable to adversarial arrangements.

Recent Field Studies Suggest, but Cannot Prove, Adversarial Allegiance

Recent studies have used exactly this strategy of examining instrument scores to explore the possibility of bias among forensic experts. Most commonly, researchers have examined scores on the Hare (2003) Psychopathy Checklist-Revised (PCL-R), a widely used instrument designed to

assess psychopathic personality features, which has become common and influential in forensic assessments of risk for violence or sexual violence (Skeem et al. 2011). Clinicians score the PCL-R on the basis of a clinical interview with an offender and a review of the offender's records. Despite some complexity and subjectivity in rating the items (Rufino et al. 2011), the PCL-R manual reports an excellent intraclass correlation rater-agreement coefficient of 0.87 for scores on the measure (Hare 2003). In other words, approximately 87% of the variability in PCL-R scores is explained by true differences in psychopathic traits, rather than any source of evaluator unreliability or measurement error. Although all tests feature some measurement error, the PCL-R is, by all conventional psychometric criteria, considered a reliable and well-developed test.

The standard error of measurement for PCL-R scores is approximately 3.0 points, which means that differences between two scores assigned to the same offender should be 6.0 points or smaller in more than 95% of cases. Differences larger than 6.0 points are unusual (less than 5% of cases in research contexts), although they do occur. If scoring differences are attributable merely to measurement error—not systematic sources of bias (such as adversarial allegiance)—the likelihood that a prosecution-retained evaluator assigns a much higher score than a defense-retained evaluator (approximately 2% of cases) should be the same as the likelihood that a defense-retained evaluator assigns a much higher score than a prosecution-retained evaluator (approximately 2% of cases). However, a pattern of findings showing that large score differences are common among evaluators, and usually reflect prosecution-retained evaluators assigning higher scores than defense-retained evaluators, would provide evidence of bias in PCL-R scoring.

In the first instrument-focused field study, researchers collected PCL-R scores assigned by prosecution-retained and defense-retained psychologists in 23 sexually violent predator cases (Murrie et al. 2008a). Sexually violent predator laws allow for the postincarceration civil commitment of sex offenders believed to be at an especially high risk for committing future sexually violent offenses (*Kansas v. Hendricks* 1997). In these trials, the difference between PCL-R scores assigned by prosecution-retained and defense-retained evaluators was large (Cohen's $d = 1.03$), leading to a low level of rater agreement ($ICC_{A,1} = 0.39$). In 61% of the cases, there was a difference larger than 6.0 points between the two PCL-R scores, although differences this great occur in fewer than 5% of cases in nonadversarial research settings (see **Table 1**). In each of these instances, it was the prosecution-retained evaluator who had assigned a higher score. For example, there was one case with a 20-point difference, with the defense-retained evaluator score falling at the 28th percentile and the prosecution-retained evaluator score falling at the 99.5th percentile.

In a follow-up study, researchers found continued evidence of bias in PCL-R scoring when they updated the sample to include 35 sexually violent predator cases (Murrie et al. 2009). They also found evidence of bias in the scores assigned on two measures designed to predict future sexual offending, the Minnesota Sex Offender Screening Tool Revised (MnSOST-R) (Epperson et al. 1998) and the Static-99 (Hanson & Thornton 2000), which evaluators score on the basis of information in offenders' correctional files. The effect size for the MnSOST-R was similar to that for the PCL-R ($d \approx 0.80$), but the effect size was somewhat smaller for the Static-99 ($d \approx 0.35$).

Similar patterns of score discrepancies that suggest adversarial allegiance have emerged in case law reviews of the PCL-R scores evaluators assigned in Canadian criminal cases (Edens et al. 2015, Lloyd et al. 2010). The size of the apparent allegiance effect varies somewhat from study to study (see **Table 1**), but in each of the available studies, scores from prosecution-retained evaluators are consistently higher and scores from defense-retained evaluators are consistently lower. Other researchers have compared scores assigned by defense and prosecution evaluators, although not necessarily to the same offenders, and found significantly higher scores from prosecution than defense evaluators ($d \approx 0.90$) in both sex offender and non-sex offender cases (DeMatteo et al. 2014a,b).

Table 1 Field study comparisons of scores assigned to the same offender by prosecution and defense evaluators

Measure/Study	Case type	Source	Sample size (N)	Effect size (d)	Differences > twice the SEM
PCL-R					
Murrie et al. (2008a)	SVP	Hearing transcripts	23	1.03	61% ^a
Murrie et al. (2009)	SVP	Hearing transcripts	35	0.83	49% ^a
Lloyd et al. (2010)	DO	Case law	15	0.58	Not reported
Edens et al. (2015)	Criminal	Case law	55	0.75	22% ^b
Static-99/Static-99R					
Murrie et al. (2009)	SVP	Hearing transcripts	27	0.37	19% ^a
MnSOST-R					
Murrie et al. (2009)	SVP	Hearing transcripts	27	0.95	52% ^a

^aPercentage of cases in which the prosecution evaluator assigned a higher score than the defense evaluator, with the difference being more than twice as large as the standard error of measurement (SEM). This type of difference should occur in only about 2% of cases by chance.

^bPercentage of cases in which the difference in scores was more than twice the SEM, but the researchers did not specify the direction of the difference. This type of difference should occur in only about 5% of cases.

Abbreviations: DO, Dangerous Offender trials for the postincarceration civil commitment of violent offenders in Canada; SVP, Sexually Violent Predator trials for the postincarceration civil commitment of sexual offenders in the United States.

Thus, even though clinicians assign very similar scores on risk measures in research and non-adversarial clinical settings, a small, but growing, body of recent research suggests prosecution-retained evaluators assign higher-risk measure scores than do defense-retained evaluators in adversarial legal settings (see **Table 1**). But are these differences simply evidence of attorney or expert selection effects? Are they attributable to the unique characteristics of the cases that proceed to trial or appeal? Or are the score differences actually caused by the adversarial arrangements underlying an expert's work? These questions can be answered only by a true experiment.

Experimental Evidence for Adversarial Allegiance

Traditionally, observational field studies are the first step in scientific research, allowing researchers to observe phenomena and generate hypotheses. But to demonstrate conclusively a phenomenon or to identify its cause, researchers must conduct strict experiments in which they can carefully control the many factors that could influence or explain the phenomenon they observed. To demonstrate the phenomenon of adversarial allegiance—as distinct from other forms of bias or selection effects—requires a situation in which attorneys are randomly assigned forensic experts, experts are randomly assigned to attorneys, experts have access to exactly the same materials, and researchers have access to data from all experts who conducted an evaluation. Only in such a scenario would discrepant scores from opposing evaluators be clearly attributable to adversarial allegiance, rather than attorney selection effects, expert selection effects, or genuine differences in the data available to opposing parties.

To explore this possibility, we recruited more than 100 practicing, doctoral-level forensic psychologists and psychiatrists and deceived them to believe they were performing a formal,

large-scale forensic consultation (Murrie et al. 2013). Unbeknownst to them, these forensic experts were randomly assigned to either a prosecution-allegiance or defense-allegiance group in which they were paid (\$400 per day) by either a “public defender service” or a “special prosecution unit.” These participants met for 10–15 minutes with an attorney who posed as leading either the public defender service or the specialized prosecution unit and requested that they score particular risk instruments on the basis of extensive offender records (a type of consultation that is common in forensic practice). Participants were led to believe that, as a group, they were reviewing and scoring cases from a large cohort. But in truth, each participant was scoring the same four case files, which spanned the range from low risk to high risk. Each set of case materials was authentic (i.e., from an actual case), including extensive records (e.g., police, court, correctional, mental health) that evaluators typically use to score risk instruments in forensic evaluations. Thus, participating forensic experts were able to score the same two commonly used risk instruments that served as the metrics for bias in earlier field studies (Murrie et al. 2008a, 2009), i.e., the Hare (2003) PCL-R and the Static-99R (Hanson & Thornton 2000, Helmus et al. 2012).¹

Overall, the risk measure scores assigned by prosecution and defense experts showed a clear pattern of adversarial allegiance. As expected, allegiance effects were stronger for the PCL-R, a measure that requires more subjective clinical judgment, than for the Static-99R, a more structured measure that permits less judgment. For the PCL-R total score, prosecution evaluators assigned significantly higher scores than did defense evaluators for three of four cases, with statistical effect sizes in the medium to large range (Cohen’s d of 0.55–0.85) and similar in magnitude to effects ($d = 0.63$ –0.83) documented in the field studies of actual sex offender trials (Murrie et al. 2009). The one case without a PCL-R allegiance effect was selected to be unusually low in psychopathy; this case received unusually low scores both from prosecution-retained and defense-retained evaluators.

In reality, courts never consider risk instrument scores that have been averaged across many experts. Rather, they usually hear expert testimony about risk scores from two experts: one called by each opposing side. To mimic this real-world arrangement as closely as possible, we performed a series of follow-up analyses to examine how likely it was that a randomly selected prosecution expert and a randomly selected defense expert would assign scores that were so different that they could not be explained by random measurement error. Findings from these comparisons revealed two clear effects. First, more than 20% of the score pairings for each case led to a score difference that was more than twice the standard error of measurement, although only approximately 4% of score pairings in research contexts led to score differences this large. For several offenders, nearly 40% of the difference scores were larger than twice the standard error of measurement, even though such score differences are quite rare in nonadversarial contexts. Second, most large (i.e., ≥ 2 times the standard error of measurement) differences were in the direction of adversarial allegiance, with the prosecution expert assigning higher scores and the defense expert assigning lower scores. This pattern was especially clear for the PCL-R, which is somewhat more lengthy and subjective than the simpler Static-99R instrument. For the three cases with clear PCL-R allegiance effects, 28% or more of all possible score pairings led to a score difference larger than twice the standard error of measurement in the direction of allegiance. Again, score differences of ≥ 2 times the standard error of measurement in one direction (e.g., prosecution > defense) should occur in only approximately 2% of cases, according to rater-agreement values from nonadversarial research contexts.

¹Only after participants had completed all scoring tasks and study procedures were they informed regarding the true study purpose and the deception on which the study was necessarily based. Though all participants had the option of withdrawing their data, none of them chose to do so. All participants still received the promised \$400 daily payment for scoring case files.

Conclusions Regarding the Effects of Adversarial Allegiance on Instrument Scores

The findings from this rigorous experiment provide strong and unambiguous evidence that even scores on ostensibly objective forensic instruments can be compromised by adversarial allegiance, at least among some experts (Murrie et al. 2013). To be clear, there was considerable variability in scores even among experts assigned to the same side, similar to the evaluator differences documented in earlier field studies (Boccaccini et al. 2008, 2014). However, neither evaluators' attitudes nor their scoring tendencies could explain the observed allegiance effects. Finding such evidence of allegiance in a strict experiment in which all other possible explanations have been excluded suggests that adversarial allegiance is indeed a significant influence among some experts. Put simply, something about working within an adversarial context caused opposing evaluators to form different opinions about the same defendant, even when the evaluators used procedures designed to enhance objectivity and reliability.

These experimental effects (Murrie et al. 2013) were somewhat smaller than effects from all but one of the field studies summarized in **Table 1**, which suggests two possibilities. First, the experimental manipulation was less powerful than experts are likely to encounter in most real cases. For example, experts in the experiment spent only 15 minutes with the retaining attorney, whereas experts in the field may have extensive contact with retaining attorneys over weeks or months (or longer if they have worked on multiple cases). In addition, experts in the experiment formed opinions based on files only, which were identical across opposing experts. But experts in the field may elicit different information by seeking different collateral sources or interviewing offenders in different ways. Furthermore, experts in the experiment were paid \$400 per day, which is not implausible but is far less than most forensic work. Therefore, the pull toward allegiance in this study was weaker compared with the pull typical of most cases in the field.

Second, in the field, selection effects probably combine with allegiance effects to produce more discrepant opinions across experts. Evaluators appear to vary in their scoring tendencies, attitudes, and opinion formation (e.g., Boccaccini et al. 2014, Miller et al. 2011, Murrie & Warren 2005, Murrie et al. 2008a), and savvy attorneys may recognize these differences and select experts strategically (or experts may self-select to the side with which they are more ideologically aligned). But once selected, experts may seek and interpret data and form opinions in a manner that is biased toward the side that retained them (Neal & Grisso 2014).

Same Score, Different Conclusion: Adversarial Allegiance and Score Interpretation

A consistent pattern across field studies and the formal experiment (Murrie et al. 2008a, 2009, 2013) was that allegiance effects were always stronger for the PCL-R (i.e., the popular measure of psychopathic personality traits, used often in legal proceedings) than for the Static-99R (i.e., the measure of sexual offender recidivism risk that is common in most sex offender proceedings). This pattern of findings likely occurs because the PCL-R is a more lengthy instrument that requires some clinician inference, whereas the Static-99R is a briefer and highly structured instrument that leaves much less room for clinician inference. So, this pattern of findings seems to suggest that allegiance effects may be problematic for only a subset of assessment instruments and could be eliminated by using only strict, objective instruments. However, findings from a recent study of Static-99R use in sex offender risk assessment cases suggest that opposing evaluators may convey different conclusions about ostensibly objective instrument results, even when they assign exactly the same score on the same instrument (Chevalier et al. 2015).

Specifically, even highly structured, objective instruments such as the Static-99R may allow some clinician discretion in interpreting the level of risk represented by a numerical score, even if there is less discretion in the actual scoring. The Static-99R manual (Phenix et al. 2012) provides separate sets of normative sample recidivism rates (“norms”) for what they label as “routine” and “high-risk” offenders. The manual also provides separate rates for 5-year and 10-year follow-up periods. In our recent survey of 109 forensic evaluators (Chevalier et al. 2015), those who routinely worked for the prosecution were much more likely to endorse score-interpretation practices indicating the highest possible level of risk (i.e., using the high risk norms), whereas those who tended to work for the defense were much more likely to endorse score-interpretation practices indicating the lowest possible level of risk (i.e., using the routine risk norms). These findings suggest that any assessment procedure that allows discretion in interpretation—no matter how objective in administration and scoring—leaves room for allegiance effects to influence the information and conclusions that experts provide to the court. Adversarial allegiance may influence score interpretation, even when it does not (or cannot) influence score assignment.

WHY DOES ADVERSARIAL ALLEGIANCE OCCUR?

Findings from field studies, surveys, and the experiment described above converge to show that performing an evaluation for one side in an adversarial setting can influence an evaluator’s decisions. We refer to this causal effect as “adversarial allegiance,” which must be separated from selection effects, evaluator differences, measurement error, and other possible causes of differences between opposing evaluators. Despite the recent surge in adversarial allegiance research, we are only beginning to understand how these factors may interact to produce differences between opposing evaluators. **Figure 1** depicts how some of these factors may compound to contribute to clear disagreements among opposing evaluators.

In any adversarial case, there is a pool of possible evaluators that an attorney may retain. These evaluators differ to some extent in their attitudes, their thresholds for coming to certain conclusions, and the typical scores they assign on forensic assessment instruments. Attorneys who are familiar with these differences can take advantage of them by selecting an expert who appears more favorable toward their side of the case.

After an attorney retains an expert, the expert’s opinion—at least in some cases—begins to favor the retaining party to a greater extent than the case data warrant. Why? None of the allegiance studies have empirically identified the precise mechanisms driving allegiance effects, but there seem to be three broad and overlapping theories on the issue. The first broad theory is relational: Adversarial allegiance may be the result of a “social-psychological process” that encourages evaluators to think of themselves as being on a side or team (Brodsky 2013, p. 161). According to this theory, adversarial allegiance is an unintentional by-product of an us-versus-them mentality inherent in adversarial contexts.

The second broad theory is that allegiance is the result of the errors in decision making that plague all areas of human judgment (Neal & Grisso 2014). This viewpoint also holds that adversarial allegiance is unintentional but attributes allegiance effects primarily to well-known, common cognitive errors in human judgment, such as favoring confirming as opposed to disconfirming evidence. Although the relational and cognitive-error explanations for adversarial allegiance effects are conceptually distinct, they can easily be viewed as complementary. The our-side mentality fostered through interactions with the retaining party may promote the development of perspectives favorable to that party, which then leads to decision-making processes that are likely to further reinforce that perspective. In short, experts may adopt the case theory of the side that retained

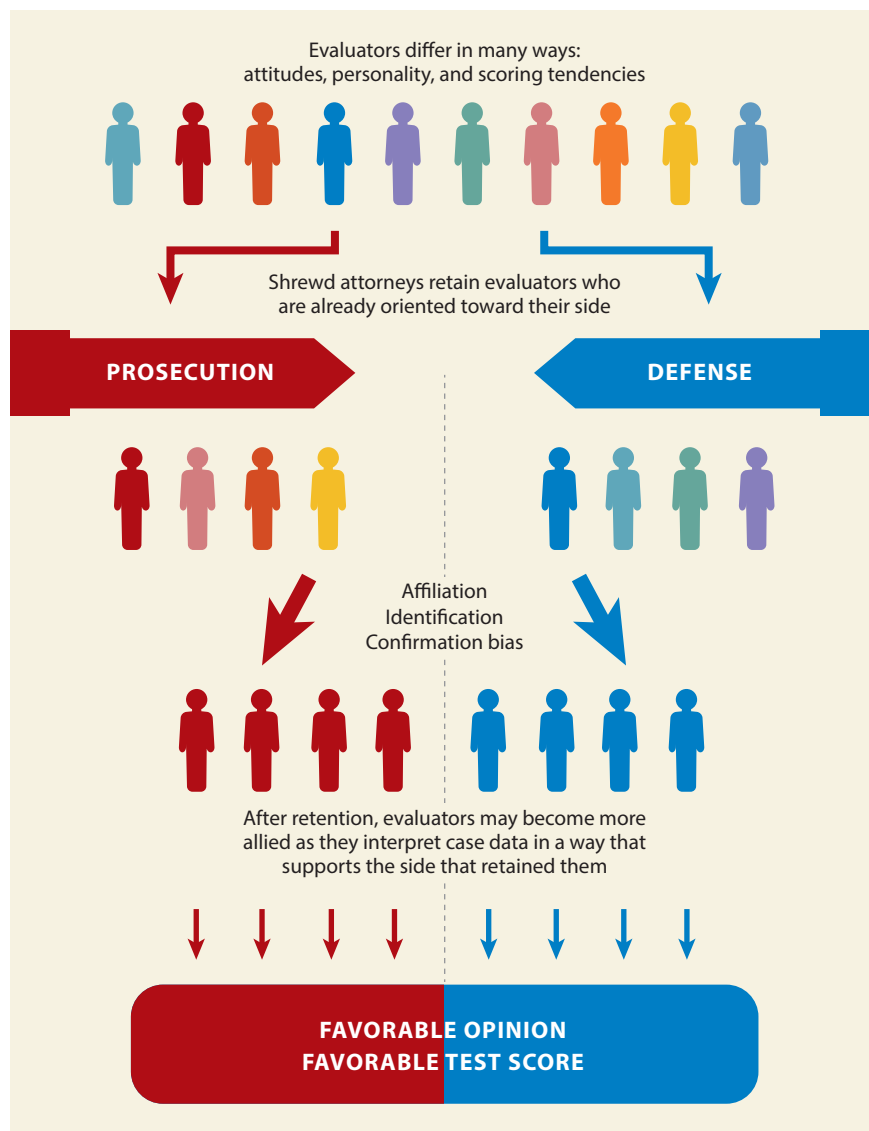


Figure 1

Evaluator differences, selection effects, and adversarial allegiance all contribute to observed differences in expert opinions.

them for relational reasons. They may then interpret data and form opinions in a manner more favorable to that perspective and less favorable to contradictory perspectives, all through processes that are probably unconscious and unintentional.

The third broad theory differs from the first two by presupposing that adversarial allegiance results from more intentional processes and motives that can probably be best summarized as financial gain (Gomez 2005, Hagen 1997). Evaluators make more money if they are called to testify in court but may be called only if their opinions favor the retaining party's case. Also,

evaluators may assume that offering favorable conclusions will increase the likelihood that the attorneys will retain them in future cases, which could also lead to adversarial allegiance effects motivated by financial interests. As one authority warns,

Like everyone else, clinicians must pay their children's college tuition, chip away at the mortgage on their house, and prepare for retirement. They know that their prospects for meeting these personal obligations depend in part on their ability to get work. . . . One of the substantial threats to clinicians' objectivity when performing forensic evaluations is the subtle pressure to meet clients' objectives lest the flow of referrals begins to wane. (Grisso 1998, p. 241)

Each of these mechanisms could have driven the allegiance effects among evaluators in recent field and experimental studies. For example, participants in the experiment (Murrie et al. 2013) met with an attorney (relational pull), were paid for their time (financial interests), and made decisions that would be expected if cognitive factors (e.g., confirmation bias) were at play. But the experiment was not designed to carefully test or disentangle these possible explanations, so the extent to which some, all, or none of these proposed mechanisms drive adversarial allegiance effects remains unknown. Furthermore, different clinicians are probably influenced to different degrees by different mechanisms.

ADVERSARIAL ALLEGIANCE IS NOT UNIQUE TO MENTAL HEALTH EXPERTS

Is adversarial allegiance unique to mental health experts or at least more common among “soft sciences” such as psychology and psychiatry in which opinions might seem more subjective? Although adversarial allegiance has been best studied in these fields, there is no reason to believe that only experts in mental health disciplines feel a pull toward allegiance. Observers have documented anecdotal evidence of adversarial allegiance in a variety of fields, ranging from medicine to accounting. To take a recent example, litigation surrounding the city of Detroit's potential bankruptcy led opposing sides to retain their own art experts to assess the value of the city's Detroit Institute for the Arts. The city's expert valued the museum's 66,000 holdings as worth \$3 to \$4 billion, whereas the creditors hired an independent expert firm that valued the holdings at more than \$8 billion (Blair 2014). Experts reviewed the same art collection, but offered different appraisals, in ways that appeared advantageous to the side retaining their services.

To consider a more widely documented example of apparent allegiance, studies and investigative reviews suggest adversarial allegiance may explain some of the problems with forensic medical evaluations of infant deaths attributed to “shaken baby syndrome” (Albert et al. 2012, Anderson 2008, Chadwick & Krous 1997). In these cases, medical experts often reach opposite conclusions when provided with the same data, to the extent that many observers had argued for reform in the rules surrounding expert testimony on this issue. In short, there appears to be no reason to believe that adversarial allegiance is unique, or even more prominent, in forensic mental health evaluations as opposed to other disciplines. Rather, allegiance effects seem to be endemic to the adversarial system, rather than to any single discipline or profession.

Recent scrutiny of subjectivity or bias across many domains of forensic science underscores the breadth of biases similar to adversarial allegiance. Recently, the US Congress mandated the National Research Council (NRC) to review broadly the state of forensic science. Their resulting publication (Comm. Identif. Needs Forensic Sci. Community, Natl. Res. Council. 2009) critically assessed a wide range of forensic science disciplines, such as analyses of DNA, hair, fibers, tool marks, bite marks, ballistics, and blood splatter. The NRC warned that the accuracy and reliability

of many popular forensic science techniques were unknown and that error rates were rarely acknowledged. Most relevant to our discussion of adversarial allegiance and bias, the NRC warned that forensic scientists are prone to a variety of contextual biases, including some that emerge because they “lack independence” from those requesting their services. For example, forensic science labs tend to work closely with—and are often located within—law enforcement and prosecutor’s offices. Thus, they may operate much as members of those teams, often receiving information about suspects and case developments that is extraneous to the analyses they perform. In other words, operating as a member of a side or team may lead to errors in interpretation and decision making, the same factors already discussed as possible mechanisms for adversarial allegiance effects.

Consistent with the NRC’s concerns, emerging research has clearly documented subjectivity and bias even in the forensic science procedures that courts have tended to consider most reliable, such as analyses of DNA (Dror & Hampikian 2011) and fingerprints (Dror & Cole 2010). This research revealing biases in forensic science procedures has proceeded in parallel with similar research addressing adversarial allegiance in forensic mental health evaluations. Although both research programs have revealed compelling evidence of bias, neither has yet developed to a point that specifically identifies, at least via controlled experiments, the precise mechanisms or processes underlying these biases. Indeed, the Committee on Identifying the Needs of the Forensic Science Community (2009) urged further research to better identify the cognitive and contextual biases that influence forensic experts.

Many forensic scientists might emphasize that their arrangements are not quite the same as those of many expert witnesses, who are retained directly by one party for one adversarial proceeding. (Likewise, mental health or medical experts who testify because they are employed by a state hospital, rather than hired for individual cases, might make a similar argument.) Thus, they may not be vulnerable to adversarial allegiance in the narrow way we have described it, but some parallels are clear. All the forensic experts who work in adversarial contexts are asked to form opinions by parties that are motivated to achieve a certain outcome. Consequently, all forensic experts inevitably approach their task with certain assumptions, expectations, and even desires. These influences—which may be described as observer effects, expectancy effects, or context effects, depending on the details (see Saks et al. 2003)—may shape forensic experts’ perceptions and interpretations of the data they are asked to consider, regardless of whether the expert has been asked to determine whether a defendant is legally insane or whether a suspect’s fingerprint matches the fingerprint smudge left at a crime scene.

Not only does adversarial allegiance seem likely across all types of expert witnesses, but some evidence also suggests it may influence nonexpert witnesses. In some early, innovative analog research, Vidmar and colleagues (Sheppard & Vidmar 1980, Vidmar & Laird 1983) assigned college undergraduates to testify as eyewitnesses in a simulated trial. They served an adversarial role, in that they were either interviewed or called to testify as a witness by either the plaintiff or the defense “lawyers,” who were also undergraduate participants. In both studies, the researchers concluded that the witnesses biased their testimony in favor of the interviewing or retaining lawyer. That is, testimony from the plaintiff’s witness seemed to support the plaintiff, and testimony from the defendant’s witness seemed to support the defense. Effects were stronger when the undergraduate “lawyers” were more Machiavellian (psychologists’ term for a self-serving personality style, so called after the infamous politician and author who espoused a ruthless approach to leadership), suggesting that shrewd lawyering can amplify the allegiance in adversarial arrangements (Sheppard & Vidmar 1980). In short, these early experiments suggest adversarial arrangements certainly exert an influence on witness testimony, whether that witness is a paid expert or not.

CONCLUSION

Adversarial allegiance among experts has long been considered a problem in the justice system, but it has only recently been a focus of empirical study. Programmatic research now reveals that the concept of adversarial allegiance is not merely a cynical theory unique to laypersons or a judiciary skeptical of expert testimony. Rather, adversarial allegiance appears to be a genuine threat to the quality of data that forensic experts provide, even when data are based on ostensibly objective instruments and procedures. Although recent research has done much to demonstrate the problem of allegiance, it has so far done little to identify and disentangle the mechanisms underlying allegiance.

Identifying these underlying mechanisms is a crucial next step in understanding allegiance and, in turn, intervening to reduce allegiance. If the causes of allegiance are primarily relational, it may be appropriate to explore interventions such as “blinding” experts to the retaining party and reducing teamlike dynamics in expert services. If causes are primarily cognitive, it may be appropriate to test procedures that minimize extraneous information and prioritize disconfirming hypotheses. If financial incentives appear to underlie problems, it may be important to explore alternate arrangements for retention and compensation. Most likely, many mechanisms are influential, though perhaps to varying degrees for different experts and across different contexts. Thus, even a best-case scenario will likely require several types of intervention that combine to reduce, but never eliminate, adversarial allegiance. At this point, we know of no empirically tested interventions to reduce allegiance among adversarial experts in any field. But we know of considerable research to suggest that the awareness of potential bias, by itself, is not sufficient to eliminate bias.

Court-appointed experts are the strategy most often proposed to remedy adversarial allegiance. However, as Mnookin (2008) emphasized, neutral experts may be problematic in those cases in which scientific data are insufficiently developed, under dispute, or open to legitimate differences of opinion. By definition, many topics of litigation are complex and open to varying perspectives, and a few may not even be amenable to a single, correct answer. In many situations, a neutral expert cannot convey to the court the nature and extent of the scientific dispute or clinical ambiguity, in ways that opposing experts may. Thus court-appointed experts may reduce the adversarial allegiance problem, but they may bring other types of bias (an empirical question that will require study) without the opportunity for scrutiny that adversarial arrangements allow.

The adversarial system has deep roots and remains deeply valued in the US judiciary. Furthermore, the judiciary will always require substantive expertise on specialized topics. Thus, as experts continue to play an important role, the challenge for research, policy, and professional guilds will be to identify ways to ensure that the forces pulling experts toward accuracy are stronger than the forces pulling them toward advocacy.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was supported in part by the National Science Foundation, Law & Social Science Program, Award SES 0961082. We thank Lucy Guarnera for her contributions to this work.

LITERATURE CITED

- Ægisdóttir S, White MJ, Spengler PM, Maugherman AS, Anderson LA, et al. 2006. The meta-analysis of clinical judgment project: fifty-six years of accumulated research on clinical versus statistical prediction. *Couns. Psychol.* 34:341–82
- Albert DM, Blanchard JW, Knox BL. 2012. Ensuring appropriate expert testimony for cases involving the “shaken baby.” *J. Am. Med. Assoc.* 308:39–40
- Am. Acad. Psychiatry Law. 2005. *Ethics Guidelines for the Practice of Forensic Psychiatry*. Bloomfield, CT: Am. Acad. Psychiatry Law. <http://www.aapl.org/ethics.htm>
- Am. Psychol. Assoc. 2013. Specialty guidelines for forensic psychology. *Am. Psychol.* 68:7–19
- Am. Psychol. Assoc. Comm. Ethical Guidel. Forensic Psychol. 1991. Specialty guidelines for forensic psychologists. *Law Hum. Behav.* 15:655–65
- Anderson M. 2008. The shaken baby debate. *Discover* 29:60–68
- Beckham JC, Annis LV, Gustafson DJ. 1989. Decision making and examiner bias in forensic expert recommendations for not guilty by reason of insanity. *Law Hum. Behav.* 13:79–87
- Bernstein DE. 2008. Expert witnesses, adversarial bias, and the (partial) failure of the Daubert revolution. *Iowa Law Rev.* 93:101–37
- Blair E. 2014. The fine art of pricing Detroit’s collection. *Natl. Public Radio, All Things Considered*, Sept 4. <http://www.npr.org/2014/09/04/345868771/the-fine-art-of-pricing-detroits-collection>
- Blair P, Marcus D, Boccaccini MT. 2008. Is there an allegiance effect for assessment instruments? Actuarial risk assessment as an exemplar. *Clin. Psychol. Sci. Pract.* 15:346–60
- Boccaccini MT, Brodsky SL. 2002. Believability of expert and lay witnesses: implications for trial consultation. *Prof. Psychol. Res. Pract.* 33(4):384–88
- Boccaccini MT, Marcus DK, Murrie DC. In press. Allegiance effects in clinical psychology research and practice. In *Psychological Science Under Scrutiny: Recent Challenges and Proposed Remedies*, ed. S Lilienfeld, ID Waldman. Hoboken, NJ: Wiley & Sons
- Boccaccini MT, Murrie DC, Rufino KA, Gardner BO. 2014. Evaluator differences in PCL-R factor and facet level scoring. *Law Hum. Behav.* 38:337–45
- Boccaccini MT, Turner D, Murrie DC. 2008. Do some evaluators report consistently higher or lower psychopathy scores than others? Findings from a statewide sample of sexually violent predator evaluations. *Psychol. Public Policy Law* 14:262–83
- Brodsky SL. 1991. *Testifying in Court: Guidelines and Maxims for the Expert Witness*. Washington, DC: Am. Psychol. Assoc.
- Brodsky SL. 2013. *Testifying in Court: Guidelines and Maxims for the Expert Witness*. Washington, DC: Am. Psychol. Assoc. 2nd ed.
- Chadwick DL, Krous HF. 1997. Irresponsible testimony by medical experts in cases involving the physical abuse and neglect of children. *Child Maltreat.* 4(2):313–21
- Chevalier C, Boccaccini MT, Murrie DC, Varela JG. 2015. Static-99R reporting practices in sexually violent predator cases: Does norm selection reflect adversarial allegiance? *Law Hum. Behav.* 39:209–18
- Comm. Identif. Needs Forensic Sci. Community, Natl. Res. Council. 2009. *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: Natl. Acad. Press
- Commons ML, Miller PM, Gutheil TG. 2012. Experts’ perceptions of expert bias. *Int. J. Law Psychiatry* 35:362–71
- Cooper J, Neuhaus IM. 2000. The “hired gun” effect: assessing the effect of pay, frequency of testifying, and credentials on the perception of expert testimony. *Law Hum. Behav.* 24(2):149–71
- DeMatteo D, Edens JF. 2006. The role and relevance of the Psychopathy Checklist-Revised in court: a case law survey of US courts (1991–2004). *Psychol. Public Policy Law* 11:62–82
- DeMatteo D, Edens JF, Galloway M, Cox J, Smith ST, et al. 2014a. Investigating the role of the Psychopathy Checklist-Revised in United States case law. *Psychol. Public Policy Law* 20:96–107
- DeMatteo D, Edens JF, Galloway M, Toney Smith S, Cox J, Koller JP. 2014b. The role and reliability of the Psychopathy Checklist-Revised in U.S. sexually violent predator evaluations: a case law survey. *Law Hum. Behav.* 38:248–55
- Diamond BL. 1959. The fallacy of the impartial expert. *Arch. Crim. Psychodyn.* 3:221–36

- Dror IE, Cole SA. 2010. The visit in “blind” justice: expert perception, judgment, and visual cognition in forensic pattern recognition. *Psychon. Bull. Rev.* 17:161–67
- Dror IE, Hampikian G. 2011. Subjectivity and bias in forensic DNA mixture interpretation. *Sci. Justice* 51:204–8
- Edens JF, Cox J, Smith ST, DeMatteo D, Sörman K. 2015. How reliable are psychopathy checklist revised scores in Canadian criminal trials? A case law review. *Psychol. Assess.* 26:447–56
- Epperson DL, Kaul JD, Goldman R, Hout S, Hesselton D, et al. 1998. *Minnesota Sex Offender Screening Tool-Revised (MnSOST-R)*. St. Paul: Minn. Dep. Correct.
- Foster WL. 1897. Expert testimony, prevalent complaints and proposed remedies. *Harvard Law Rev.* 11:169–86
- Franklin K. 2012. Forensic psychologist blackballed over competency opinions. In *the News: Forensic Psychol. Criminol. Psychol.-Law*, Feb. 28. <http://forensicpsychologist.blogspot.com/2012/02/forensic-psychologist-blackballed-over.html>
- Goldstein AG. 2006. Forensic psychology: toward a standard of care. In *Forensic Psychology: Emerging Topics and Expanding Roles*, ed. A Goldstein, pp. 683–707. Hoboken, NJ: Wiley & Sons
- Gomez JCB. 2005. Silencing hired guns: ensuring honesty in medical expert testimony through state legislation. *J. Legal Med.* 26:385–99
- Grisso T. 1998. *Forensic Evaluation of Juveniles*. Sarasota, FL: Prof. Res. Press
- Grisso T. 2003. *Evaluating Competencies: Forensic Assessments and Instruments*. New York: Plenum
- Grove WM, Zald DH, Lebow BS, Snitz BE, Nelson C. 2000. Clinical versus mechanical prediction: a meta-analysis. *Psychol. Assess.* 12:19–30
- Hagen MA. 1997. *Whores of the Court: The Fraud of Psychiatric Testimony and the Rape of American Justice*. New York: Regan
- Hand L. 1901. Historical and practical considerations regarding expert testimony. *Harvard Law Rev.* 15:40–58
- Hanson RK, Morton-Bourgon KE. 2009. The accuracy of recidivism risk assessments for sexual offenders: a meta-analysis of 118 prediction studies. *Psychol. Assess.* 21:1–21
- Hanson RK, Thornton D. 2000. Improving risk assessments for sex offenders: a comparison of three actuarial scales. *Law Hum. Behav.* 24:119–36
- Hare RD. 2003. *The Hare PCL-R: Technical Manual*. Toronto, Can.: Multi-Health Syst. 2nd ed.
- Harris PB, Boccaccini MT, Murrie DC. 2014. Rater differences in psychopathy measure scoring and predictive validity. *Law Hum. Behav.* doi:10.1037/lhb0000115
- Heilbrun K. 2009. *Evaluation for Risk of Violence in Adults*. New York: Oxford Univ. Press
- Heilbrun K, Brooks S. 2010. Forensic psychology and forensic science: a proposed agenda for the next decade. *Psychol. Public Policy Law* 16(3):219–53
- Helmus L, Thornton D, Hanson RK, Babchishin KM. 2012. Improving the predictive accuracy of Static-99 and Static-2002 with older sex offenders: revised age weights. *Sex. Abuse J. Res. Treat.* 24:64–101
- Huber PW. 1993. *Galileo’s Revenge: Junk Science in the Courtroom*. New York: Basic Books
- Kansas v. Hendricks*, 521 U.S. 346 (1997)
- Kassin SM, Drizin SA, Grisso T, Gudjonsson GH, Leo RA, et al. 2009. Police-induced confessions: risk factors and recommendations. *Law Hum. Behav.* 34:3–38
- Krafka C, Dunn MA, Johnson MT, Cecil JS, Miletich D. 2002. Judge and attorney experiences, practices, and concerns regarding expert testimony in federal civil trials. *Psychol. Public Policy Law* 8:309–32
- Lloyd CD, Clark HJ, Forth AE. 2010. Psychopathy, expert testimony, and indeterminate sentences: exploring the relationship between Psychopathy Checklist-Revised testimony and trial outcome in Canada. *Legal Criminol. Psychol.* 15(2):323–39
- Luborsky L, Diguer L, Seligman DA, Rosenthal R, Krause ED, et al. 1999. The researcher’s own therapy allegiances: a “wild card” in comparisons of treatment efficacy. *Clin. Psychol. Sci. Pract.* 6:95–106
- Luborsky L, Singer B, Luborsky L. 1975. Comparative studies of psychotherapies: Is it true that everyone has won and all must have prizes? *Arch. Gen. Psychiatry* 32(8):995–1008
- McDermott PA, Watkins MW, Rhoad AM. 2014. Whose IQ is it? Assessor bias variance in high-stakes psychological assessment. *Psychol. Assess.* 26:207–14

- Miller AK, Rufino KA, Boccaccini MT, Jackson RL, Murrie DC. 2011. On individual differences in person perception: raters' personality traits relate to their Psychopathy Checklist-Revised scoring tendencies. *Assessment* 18:253–60
- Mnookin J. 2007. Idealizing science and demonizing experts: an intellectual history of expert evidence. *Va. Law Rev.* 52(101):8–11
- Mnookin J. 2008. Expert evidence, partisanship, and epistemic confidence. *Brooklyn Law Rev.* 73:587–611
- Mossman D. 1999. "Hired guns," "whores," and "prostitutes": case law references to clinicians of ill repute. *J. Am. Acad. Psychiatry Law* 27(3):414–25
- Murrie DC, Boccaccini MT, Guarnera LA, Rufino KA. 2013. Are forensic experts biased by the side that retained them? *Psychol. Sci.* 24:1889–97
- Murrie DC, Boccaccini MT, Johnson JT, Janke C. 2008a. Does interrater (dis)agreement on Psychopathy Checklist scores in sexually violent predator trials suggest partisan allegiance in forensic evaluations? *Law Hum. Behav.* 32(4):352–62
- Murrie DC, Boccaccini MT, Turner DB, Meeks M, Woods C, et al. 2009. Rater (dis)agreement on risk assessment measures in sexually violent predator proceedings: evidence of adversarial allegiance in forensic evaluation? *Psychol. Public Policy Law* 15(1):19–53
- Murrie DC, Boccaccini MT, Zapt PA, Warren JI, Henderson CE. 2008b. Clinician variation in findings of competence to stand trial. *Psychol. Public Policy Law* 14:177–93
- Murrie DC, Warren JI. 2005. Clinician variation in rates of legal sanity opinions: implications for self-monitoring. *Prof. Psychol. Res. Pract.* 36:519–24
- Neal TMS. 2011. *The objectivity demand: experiences and behaviors of psychologists in capital case evaluations*. PhD Thesis, Univ. Alabama, Tuscaloosa
- Neal TMS, Grisso T. 2014. The cognitive underpinnings of bias in forensic mental health evaluations. *Psychol. Public Policy Law* 20:200–11
- Otto RK. 1989. Bias and expert testimony of mental health professionals in adversarial proceedings: a preliminary investigation. *Behav. Sci. Law* 7(2):267–73
- Otto RK, DeMier R, Boccaccini MT. 2014. *Forensic Reports and Testimony: Effective Communication for Psychologists and Psychiatrists*. Hoboken, NJ: Wiley
- Otto RK, Douglas KS, eds. 2010. *Handbook of Violence Risk Assessment*. New York: Routledge
- Otto RK, Heilbrun K. 2002. The practice of forensic psychology: a look toward the future in light of the past. *Am. Psychol.* 57:5–18
- Phenix A, Helmus L, Hanson RK. 2012. *Static-99R and Static-2002R Evaluators' Workbook*, July. <http://static99.org/>
- Rogers R. 1987. Ethical dilemmas in forensic evaluations. *Behav. Sci. Law* 5:149–60
- Rogers R, Fiduccia CE. 2015. Forensic assessment instruments. In *APA Handbook of Forensic Psychology: Volume 1. Individual and Situational Influences in Criminal and Civil Contexts*, ed. BL Cutler, PA Zapf, pp. 19–34. Washington, DC: Am. Psychol. Assoc.
- Rufino KA, Boccaccini MT, Guy LS. 2011. Scoring subjectivity and item performance on measures used to assess violence risk: the PCL-R and HCR-20 as exemplars. *Assessment* 18:453–63
- Saks MJ, Risinger DM, Rosenthal R, Thompson WC. 2003. Context effects in forensic science: a review and application of the science of science to crime laboratory practice in the United States. *Sci. Justice* 43(2):77–90
- Sheppard BE, Vidmar N. 1980. Adversary pretrial procedures and testimonial evidence: effects of lawyer's role and Machiavellianism. *J. Personal. Soc. Psychol.* 39:320–32
- Shuman DW, Whitaker E, Champagne A. 1994. An empirical examination of the use of expert witnesses in the courts—part II: a three city study. *Jurimetrics* 34:193–208
- Skeem JL, Polaschek DL, Patrick CJ, Lilienfeld SO. 2011. Psychopathic personality bridging the gap between scientific evidence and public policy. *Psychol. Sci. Public Interest* 12(3):95–162
- United States v. Alvarez*, 519 F. 2d 1036 (3rd Cir. 1975)
- United States ex rel. Edney v. Smith*, 425 F. Supp. 1038 (E.D.N.Y. 1976)
- Vidmar N, Laird NM. 1983. Adversary social roles: their effects on witnesses' communication of evidence and the assessment of adjudicators. *J. Personal. Soc. Psychol.* 44:888–98

- Vitacco MJ, Erikson SK, Kurus S, Apple BN. 2012. The role of the Violence Risk Appraisal Guide and Historical, Clinical, Risk-20 in U.S. courts: a case law survey. *Psychol. Public Policy Law* 18:361–91
- Wells GL, Small M, Penrod S, Malpass RS, Fulero SM, Brimacombe CAE. 1998. Eyewitness identification procedures: recommendations for lineups and photospreads. *Law Hum. Behav.* 22:1–39
- Wigmore J. 1923. *A Treatise on the Anglo-American System of Evidence in Trials at Common Law: Including the Statutes and Judicial Decisions of All Jurisdictions of the United States and Canada*. Boston, MA: Little, Brown
- Zusman J, Simon J. 1983. Differences in repeated psychiatric examinations of litigants to a lawsuit. *Am. J. Psychiatry* 140:1300–4



Contents

An Unusual Career: Considering Political/Legal Orders and Unofficial Parallel Realities <i>Sally Falk Moore</i>	1
Housing, Poverty, and the Law <i>Matthew Desmond and Monica Bell</i>	15
Adversarial Allegiance among Expert Witnesses <i>Daniel C. Murrie and Marcus T. Boccaccini</i>	37
Approaching the Econo-Socio-Legal <i>Amanda Perry-Kessaris</i>	57
Behavioral Ethics Meets Legal Ethics <i>Jennifer K. Robbennolt</i>	75
Constitutional Amendments <i>Heinz Klug</i>	95
Effectiveness and Social Costs of Public Area Surveillance for Crime Prevention <i>Brandon C. Welsh, David P. Farrington, and Sema A. Taberi</i>	111
Empirical Comparative Law <i>Holger Spamann</i>	131
Experimental Justice Reform: Lessons from the World Bank and Beyond <i>Deval Desai and Michael Woolcock</i>	155
Getting Around and Getting On: Self-Interested Resistance to Technology in Law Enforcement Contexts <i>Helen Wells</i>	175
Governing Global Supply Chains: What We Know (and Don't) About Improving Labor Rights and Working Conditions <i>Daniel Berliner, Anne Regan Greenleaf, Milli Lake, Margaret Levi, and Jennifer Noveck</i>	193

Improving the Effectiveness of Suspect Interrogations <i>Christian A. Meissner, Christopher E. Kelly, and Skye A. Woestehoff</i>	211
Legal Control of Marginal Groups <i>Forrest Stuart, Amada Armenta, and Melissa Osborne</i>	235
Misdemeanors <i>Alexandra Natapoff</i>	255
Race, Ethnicity, and Culture in Jury Decision Making <i>Jennifer S. Hunt</i>	269
Social Rights Constitutionalism: Negotiating the Tension Between the Universal and the Particular <i>Daniel M. Brinks, Varun Gauri, and Kyle Shen</i>	289
South Asian Legal History <i>Mitra Sbarafi</i>	309
Surveying Key Aspects of Sociolegal Scholarship on India: An Overview <i>Jayanth K. Krishnan and Patrick W. Thomas</i>	337
The Economics of Civil Procedure <i>Daniel Klerman</i>	353
The Fall and Rise of Law and Social Science in China <i>Sida Liu and Zhizhou Wang</i>	373
The Relevance of Law to Sovereign Debt <i>W. Mark C. Weidemaier and Mitu Gulati</i>	395
Indexes	
Cumulative Index of Contributing Authors, Volumes 1–11	409
Cumulative Index of Article Titles, Volumes 1–11	412

Errata

An online log of corrections to *Annual Review of Law and Social Science* articles may be found at <http://www.annualreviews.org/errata/lawsocsci>